

VERIFICATION FOR TRANSLATION

I, Ryoko MIDORIKAWA, a national of Japan, c/o SAEKI & PARTNERS of 4th Floor, Aminosan Kaikan Building, 15-8, Nihonbashi 3-chome, Chuo-ku, Tokyo 103-0027, Japan, do hereby solemnly and sincerely declare:

1) THAT I am well acquainted with the Japanese language, English language, and

2) THAT the attached is a true, accurated and faithful translation into the English language made by me of Japanese Patent Application No. 2000-133519 filed to the Japanese Patent Office on May 2, 2000.

Signed this 22 th day of August , 2005.

A handwritten signature in cursive script, reading "Ryoko Midorikawa", is written over a horizontal line.

Ryoko MIDORIKAWA

[Document] Specification

[Title of the Invention] Novel Nucleic Acid Base Pair

[Claims]

[Claim 1] A method for constructing selective base pair comprising introducing a group having ability to form steric hindrance and electrostatic repulsion as well as stacking action between bases in nucleic acid base.

[Claim 2] The method according to claim 1 wherein the group having ability to form steric hindrance, electrostatic repulsion and stacking action is a group to hinder formation of base pair with base part of natural nucleic acid.

[Claim 3] The method according to claim 1 or claim 2 wherein the group having ability to form steric hindrance and electrostatic repulsion, and stacking action is aromatic heterocyclic group.

[Claim 4] The method according to claim 3 wherein aromatic heterocyclic group is five membered or six membered aromatic heterocyclic group having 1 or 2 sulfur atoms, oxygen atoms or nitrogen atoms as a heteroatom.

[Claim 5] The method according to claim 4 wherein aromatic heterocyclic group is thiophene.

[Claim 6] The method according to any of claims 1 - 5 comprising introducing a group to be able to form additional hydrogen bonds.

[Claim 7] The method according to any of claims 1 - 6 wherein the base pair is a base pair which can be recognized by polymerase.

[Claim 8] The method according to claim 7 wherein the polymerase is DNA polymerase or RNA polymerase.

[Claim 9] A method for designing nucleic acid to construct selective

base pair comprising utilizing steric hindrance and electrostatic repulsion, and stacking action in the nucleic acid base part.

[Claim 10] A method for designing nucleic acid to construct selective base pair comprising hindering to construct base pair with the natural nucleic acid base part by utilizing steric hindrance and electrostatic repulsion and stabilizing to impart stacking action.

[Claim 11] The method for designing nucleic acid according to claim 9 or claim 10 wherein the nucleic acid base pair is a base pair which can be recognized by polymerase.

[Claim 12] A nucleic acid comprising being designed by the method according to any of claims 9 - 11.

[Claim 13] The nucleic acid according to claim 12 wherein the nucleic acid has base comprising purine derivatives having a group at position-6 which can generate steric hindrance.

[Claim 14] The nucleic acid according to claim 13 wherein the nucleic acid base is 2-amino-6-thienyl-purine or derivatives thereof.

[Claim 15] The nucleic acid according to claim 12 wherein the nucleic acid has base containing pyridine having hydroxyl group or keto group at position-2.

[Claim 16] The nucleic acid according to claim 15 wherein the nucleic acid base is pyridine-2-one or tautomer thereof.

[Claim 17] The nucleic acid according to any of claims 12 - 16 wherein the nucleic acid is a nucleic acid constructing base pair with its complementary nucleic acid.

[Claim 18] A process for production of the nucleic acid comprising the

nucleic acid according to any of claims 12 - 17.

[Claim 19] The process according to claim 18 wherein the nucleic acid is another nucleic acid to construct base pair.

[Claim 20] A codon comprising one or more nucleic acid according to any of claims 12 - 17.

[Claim 21] The codon according to claim 20 wherein the codon encodes amino acids.

[Claim 22] The codon according to claim 21 wherein amino acids are non-natural amino acids.

[Claim 23] A nucleic acid molecule comprising containing the nucleic acid according to any of claims 12 - 17 and natural nucleic acid.

[Claim 24] The nucleic acid molecule according to claim 23 wherein the nucleic acid molecule encodes proteins.

[Claim 25] The nucleic acid molecule according to claim 23 or claim 24 wherein the nucleic acid molecule has whole or part of genetic information of the natural gene.

[Claim 26] A process for production of nucleic acid having complementary strand thereof comprising reacting the polymerase with the nucleic acid according to any of claims 23 - 25.

[Claim 27] The method according to claim 26 wherein the polymerase is DNA polymerase or RNA polymerase.

[Claim 28] A process for production of non-natural nucleic acid comprising inserting or substituting one or more nucleic acid according to any of claims 12 - 17 in the natural nucleic acid.

[Claim 29] The process for production of non-natural nucleic acid

according to claim 28 wherein a position, to which the nucleic acid according to any of claims 12 - 17 is inserted or substituted, has a codon unit and the other part has base sequence encoding natural amino acid sequence.

[Claim 30] A process for production of protein having amino acid sequence based on codons of the nucleic acid according to any of claims 23 -25 or the non-natural nucleic acid obtained by the method according to claim 28 or claim 29.

[Claim 31] The process for production of protein according to claim 30 comprising being inserted or substituted by the non-natural amino acid in the part or whole of amino acid sequence of natural protein.

[Claim 32] A microorganism which is transformed by non-natural gene which can be produced by the process according to claim 28 or claim 29.

[Claim 33] A method for screening functions of amino acids coded by natural gene comprising using the non-natural gene which can be produced by the method according to claim 28 or claim 29.

[Detailed Description of Invention]

[0001]

[Field of the Invention]

The present invention relates to formation of selective novel artificial nucleic acid base pair by utilizing steric hindrance.

The present invention further relates to replication and transcription of nucleic acid using the novel artificial nucleic acid base pair of the present invention, and a system for protein synthesis or functional nucleic acid. More particularly, the present invention pertains the novel artificial nucleic acid having properties to form selective base pair by applying steric hindrance,

electrostatic repulsive force and stacking action, a process for production thereof, codon containing the same, nucleic acid molecule containing the same, a process for production of novel protein using the above nucleic acid molecules or non-natural gene.

[0002]

[Related Art]

Genetic information of organisms in the earth are transferred by using nucleic acids comprising of four bases consisting of adenine (A), guanine (G), cytosine (C) and thymine (T) as a gene. Proteins are synthesized according to genetic informations of mRNA which is transcribed from DNA of gene. In that occasion, 64 types of codon consisting of 3 bases ($4^3 = 64$) correspond to 20 types of amino acids.

If novel nucleic acid base (X and Y, in which X and Y form specific base pair) can be created in addition to already known for bases (A, G, C, T), numbers of codon can be increased greatly ($6^3 = 216$). As a result, proteins containing non-natural amino acids can possibly be synthesized by matching the newly created codons with non-natural amino acids [J. D. Bain, et al. Nature, 356, 537-539 (1992)].

[0003]

Heretofore, a pair of isocytosine and isoguanine has been reported as an artificial base pair except for A-T and G-C. Isoguanine tends to form base pair with thymine due to tautomerism of isoguanine [C. Switzer, et al. J. Am. Chem. Soc. 111, 8322-8323 (1989); C. Y. Switzer, et al. Biochemistry 32, 10489-10496 (1993)]. Several novel base pairs have been reported, but there were problems on recognition by polymerase and no practical use has known

[J. A. Piccirilli, et al., Nature, 343, 33-37 (1990); J. Horlacher, et al. Proc. Natl. Acad. Sci. USA, 92, 6329-6333 (1995); J. C. Morales, et al., Nature struct. biol., 5, 954-959 (1998)].

[0004]

Nucleic acid molecules having various functions were found by in vitro selection method [A. D. Ellington, et al. Nature 346, 818-822 (1990); C. Tuerk, et al. Science 249, 505-510 (1990)]. If the novel base pair X-Y hereinabove can be recognized by polymerases such as DNA polymerase, RNA polymerase and reverse transcriptase, the present in vitro selection method using 4 bases can be performed by using 6 bases, then possibility to create nucleic acid molecules having novel function, which could not be practically realized by using 4 bases, can be expected.

Further, creation of novel base pair has expected for treatment of hereditary diseases caused by gene abnormality, in which one or more base in the gene is replaced by different base.

[0005]

In order to form selective base pair with non-natural bases, the present inventors have developed a method to hinder base pairing with non-natural base and natural base by utilizing steric hindrance. They have signified that nucleotide substrate of corresponding non-natural pyridine-2-one (dyTP and ryTP) is incorporated selectively into DNA and RNA with DNA polymerase and RNA polymerase by employing template DNA containing 2-amino-6-(N,N-dimethylamino)-9-(2-deoxy- β -D-ribofuranosyl) purine (dx1) having non-natural base, and have already filed patent application (Japanese Patent Application No.

11-201450/1999) . These are shown in Fig.1 a and b.

[0006]

This compound could hinder base pairing with natural base such as thymine (or uridine) (refer to Fig. 1 b) and cytosine in some extent due to steric bulkiness of dimethylamino group in the base (dx1 in Fig. 1 a). However, this steric hindrance could affect to the neighboring bases, and simultaneously could give inferior effect on stacking between bases, and resulted low rate of incorporation of dyTP by Klenow fragment as well as incapable of suppressing incorporation of thymidine triphosphate (dTTP) to dx1.

[0007]

[Problems to be Resolved by the Invention]

The present invention provides ideas for selective formation of novel artificial nucleic acid base pair as a result of recognition of base pairing by polymerase such as DNA polymerase by utilizing steric hindrance between base pairs, further generating steric hindrance between only base pair plane without giving deterioration for stacking between bases and more preferably by selecting bases utilizing electrostatic repulsion against natural bases.

An aspect of the present invention is to provide novel artificial nucleic acid base pair which does not form base pair with natural nucleic acid and forms selective base pair in themselves as well as being recognized by various polymerases. Further aspect of the present invention is to provide artificial nucleic acid, codon containing the same, nucleic acid molecule, non-natural gene and application thereof.

[0008]

[Means for Solving the Problems]

The present inventors have studied extensively to create novel artificial nucleic acid base pairs, which could not form base pair with natural nucleic acid, but could selectively form base pair by themselves and could be recognized by various polymerases, and have found out that formation of nucleic acid base pair with natural nucleic acid could be inhibited by applying steric hindrance of nucleic acid base, and formation of selective base pair between newly designed nucleic acid bases could be made. Further, it was found out that such newly designed nucleic acids could be recognized by various natural polymerases.

For example, in order not to form base pair with thymine but to form steric hindrance with keto group at position-6 of thymine, 2-amino-6-(N,N-dimethylamino) purine (hereinafter designates as X1), in which two bulky methyl groups are introduced in amino group at position-6 of 2,6-diaminopurine, is designed. As a result, the X1 does not form base pair with thymine, but bases such as pyridine-2-one (hereinafter designates as Y), an analog of thymine, in which oxo group at position-6 is replaced by hydrogen atom, can form base pair with X1 (refer to Fig. 1 a and b) and thus patent application was filed (Japanese Patent Application No. 11-201450/1999).

It was able to hinder base pairing with natural base such as thymine (or uridine) (refer to Fig. 1 b) and cytosine to some extent merely by steric hindrance of nucleic acid base, however, it was not able to obtain sufficient selectivity.

[0009]

The present inventors have examined novel artificial base pair by considering not only steric hindrance but also electrostatic repulsion between bases and stacking action with the neighboring bases, and could obtain artificial base pair with superior selectivity.

[0010]

The present invention provides novel artificial nucleic acid base pair capable of forming selective base pair by utilizing steric hindrance which can be recognized by polymerases such as DNA polymerase, and novel artificial genes.

[0011]

The present invention relates to a method for constructing selective base pair comprising introducing a group having ability to form steric hindrance, electrostatic repulsion and stacking action in nucleic acid base. More particularly, the present invention relates to a method for constructing selective base pair to hinder formation of base pair with base part of natural nucleic acid by an action of said steric hindrance and electrostatic repulsion, and to form stable structure with neighboring bases by the stacking action, and the said base pair can be recognized by polymerase.

Further, the present invention relates to a method for designing nucleic acid to construct selective base pair comprising utilizing steric hindrance and electrostatic repulsion, and stacking action in the nucleic acid base part. More particularly, the present invention relates to a method for designing nucleic acid to construct selective base pair comprising hindering to construct base pair with the natural nucleic acid base part by utilizing steric hindrance and electrostatic repulsion and stabilizing with the neighboring

bases by the stacking action, and the said base pair can be recognized by polymerase.

The present invention relates to a nucleic acid, which can construct selective base pair, prepared by introducing a group having ability to form steric hindrance and electrostatic repulsion, and stacking action in nucleic acid base. More particularly, the present invention relates to a nucleic acid for constructing selective base pair wherein the said group having ability to form steric hindrance and electrostatic repulsion is to hinder formation of base pair with base part of natural nucleic acid, and more preferably to stabilize with the neighboring bases by the stacking action, and the said base pair can be recognized by polymerase.

[0012]

The present invention discloses novel artificial nucleic acid which have similar behavior with nucleic acids containing natural bases and a method for designing such the nucleic acid. The nucleic acid of the present invention can be applied in the similar manner as the natural nucleic acid.

Consequently, the present invention relates to various applications using the nucleic acid of the present invention or nucleic acid designed by the method of the present invention.

More particularly, the present invention relates to a codon comprising one or more nucleic acid designed by the nucleic acid of the present invention or nucleic acid designed by the method of the present invention. The said codon can encode amino acids in the similar manner as the natural nucleic acid. The said amino acids can be non-natural amino acids. Further, the present invention relates to a nucleic acid molecule containing the nucleic

acid of the present invention, the nucleic acid designed by the method of the present invention or the nucleic acid of the natural origin. The said nucleic acid molecule can encode proteins in the similar manner as the natural nucleic acid. Further the said nucleic acid molecule can maintain whole or part of genetic informations of the natural gene. Nucleic acid having complementary strand can be prepared by an action of various polymerases on such the nucleic acid molecule. The present invention also relates to such the process for production of complementary strands.

[0013]

In addition, the nucleic acid of the present invention or the nucleic acid designed by the method of the present invention can be introduced or substituted to a part of natural gene. Consequently, the present invention relates to a process for production of non-natural gene comprising introducing or substituting one or more nucleic acid of the present invention or the nucleic acid designed by the method of the present invention into the natural gene. The introduction or substitution can be performed with the codon unit of the present invention as described hereinbefore.

Further, the present invention relates to a process for production of protein having amino acid sequence based on codons of the non-natural gene or the nucleic acid of the present invention. Protein to which non-natural amino acid is introduced or substituted in the part of natural protein can be produced in case that codon containing the nucleic acid of the present invention or the nucleic acid designed by the method of the present invention encodes non-natural amino acid.

[0014]

Consequently, the present invention provides a process for production of novel protein comprising substituting or introducing other natural or non-natural amino acid, preferably non-natural amino acid in a part of natural protein by the method of the present invention. According to this method, functions of amino acids in the protein coded by natural gene can be screened. The present invention also relates to a method for screening function of each amino acid of protein encoded by natural gene.

The present invention also relates to a microorganism transformed by non-natural gene containing the nucleic acid of the present invention or nucleic acid designed by the method of the present invention (hereinafter simply designates as the nucleic acid of the present invention).

Further, since the novel base pair of the present invention does not constitute base pairing with the natural bases, it is useful for treatment of hereditary diseases caused by gene in which one or more base is replaced by other base. The present invention provides pharmaceutical composition comprising novel base pair or a base in the said base pair.

[0015]

An object of the present invention is to provide artificial nucleic acid which does not form base pair with a base of the natural nucleic acid and can be recognized by polymerase. The conventional artificial nucleic acid has produced by attempting to change at the position of hydrogen bond, as a result, base pairing with a base of the natural nucleic acid could not be hindered substantially as well as showing insufficient base pair selectivity. The present inventors have solved such a problem by introducing a group forming steric hindrance and electrostatic repulsion and having stacking

action. The present invention provides novel artificial nucleic acid which can form selective base pairing with artificial nucleic acids themselves.

Consequently, the present invention will be explained more concretely by referring examples hereinbelow, but these examples are illustrated only for the purpose of better understanding of the present invention, and the fact that the present invention is not limited by these examples is obvious according to the technical idea of the present invention explained hereinbefore.

[0016]

Examples of groups having actions for steric hindrance, electrostatic repulsion and stacking action on base part of the nucleic acid are group having steric hindrance which hinders hydrogen bonding having deteriorated action between bases, having electrostatic repulsive force and having π electron for stacking action. These groups are not limited if they have deteriorating actions as a nucleic acid base. More preferably, a group having size not to hinder base pairing for other nucleic acid is preferable. Further, groups without having polar site for hydrogen bond and activated hydrogen atom are preferable, however if these polar site or activated hydrogen is located at distal position where hydrogen bonding may be impossible, it may not necessary to consider.

Examples of groups having steric hindrance, electrostatic repulsion and stacking action in the base pair of the present invention are preferably aromatic heterocyclic group having planar structure. Such aromatic heterocyclic group has sufficient size for steric hindrance on the planar direction of molecule, and can generate electrostatic repulsion by different

atoms, and also is expected to show stacking action by π electron of aromatic heterocyclic group.

[0017]

Examples of such the aromatic heterocyclic group are, concretely, five membered or six membered aromatic heterocyclic group having one or two sulfur atom, oxygen atom or nitrogen atom. These aromatic heterocyclic groups can be condensed-ring, polycyclic or monocyclic group. Among them, monocyclic group is preferable due to steric size. These aromatic heterocyclic groups can have any substituents, but a group without having large substituent is preferable due to possibility to cause stereospecific limitation or generation of deteriorative hydrogen bonding. Examples of substituents are hydroxyl, amino, carbonyl, lower alkyl of carbon 1 - 5, lower alkoxy, lower alkylamino or nitro.

[0018]

Conventional chemical synthesis can be applied for methods of introducing groups to form steric hindrance, electrostatic repulsion or stacking action in the base.

[0019]

Nucleic acid of the present invention is artificial nucleic acid which can be recognized by polymerase. Examples of polymerase can be any polymerase, preferably DNA polymerase and RNA polymerase. Recent studies on structural analysis of polymerase indicates that interaction of polymerase and nucleic acid is essentially identical with each other. Formation of base pair of the present invention relates to essential nature of polymerase, consequently, the base pair formation of the present invention can be utilized

not only for DNA polymerase and RNA polymerase but also for all polymerase including reverse transcriptase.

[0020]

Further, configuration of nucleic acid can be calculated by analysis of molecular configuration or precise determination of distance between atoms. Consequently, by applying these results, chemical structure, which causes steric hindrance in one side and provides one or more hydrogen bonds, preferably two hydrogen bonds in other side, can be designed. Consequently, the present invention includes a method for designing artificial nucleic acid based on steric hindrance of the nucleic acid, preferably steric hindrance in the base part of nucleic acid, electrostatic repulsion and stacking action. In the designing base pair of the present invention, designing based on Watson-Crick type base pair is conventional, but Hoogsteen base pairing may also be applicable.

[0021]

Nucleic acid of the present invention can be a nucleic acid designed by steric hindrance of nucleic acid, preferably designed by steric hindrance in the base part of the nucleic acid, and is preferable to form selective base pairing with each of artificial nucleic acid. Preferably, base pairing of artificial nucleic acid can be recognized by polymerase and more preferably the complementary strand can be constructed similar to the natural nucleic acid by and action of polymerase.

Nucleic acid of the present invention can be synthesized by conventional chemical synthesis but is not limited to that method. Chemical synthesis is exemplified in Fig. 2.

[0022]

Method for incorporating nucleic acid of the present invention into the nucleic acid sequence can be performed by applying conventional method for incorporation of natural nucleic acid or by applying similar method thereof. For example, a method using DNA synthesizer, method for using polymerase and point mutation technology can be mentioned. Labeling can also be possible made as same as in the natural nucleic acid.

The present invention also includes nucleic acid which can be used for gene fragment or probe, and include nucleic acid molecule containing the nucleic acid of the present invention. The nucleic acid molecule of the present invention contains one or more nucleic acid of the present invention, and can be a single strand or double strands. Non-natural gene of the present invention includes natural gene in which whole of part of it is replaced by nucleic acid of the present invention, natural gene to which one or more nucleic acid of the present invention is added, or combination thereof. Such the non-natural gene of the present invention can be modified by the same or similar method used for the conventional modification of natural gene.

[0023]

Consequently, nucleic acid molecule or non-natural gene of the present invention can be used for transformation of microorganisms by the same way as in the conventional natural gene by inserting suitable vector or phage and inserted into microorganisms to produce transformant containing the artificial nucleic acid of the present invention.

[0024]

Further, new codon containing nucleic acid of the present invention

can be designed. For example, the present novel artificial nucleic acid base is set as X and Y, combination thereof such as XXY, XYX, YXX, a combination by themselves, and AXA, TYT, CGX, ATX, which are combination of base of natural nucleic acid and artificial base of the present invention. Such codons can be designed. New codons can code natural amino acid, or non-natural amino acid. Further, functions such as transcription, transfer can be coded. Accordingly, the present invention not only provide novel artificial nucleic acid, but also providing possibility of designing completely new genetic code by designing new codon containing nucleic acid of the present invention.

[0025]

As a result of designing t-RNA corresponding to new codon of the present invention. New protein synthesis system can be designed by which large number of amino acid can be utilized.

Usable amino acid can be amino acid utilized on protein synthesis on liposome. Consequently, the present invention provides novel protein synthetic system using codon of the present invention.

Heretofore, some amino acids in the natural protein are very difficult to substitute non-natural amino acid, or insertion of non-natural amino acid into the natural protein is also very difficult. According to the protein synthesis system of the present invention, proteins containing desired non-natural amino acid can be produced by substituting or inserting the nucleic acid having codon of desired position into the nucleic acid of the present invention. And such the conversion of amino acid resulted to make screening functions of amino acid in the protein.

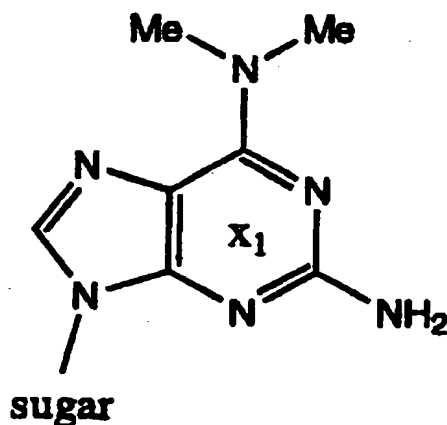
[0026]

The present invention will be explained by the following examples in detail.

In order not to form base pairing 2,6-diaminopurine with thymine, two bulky methyl groups were introduced into the amino group at position-6 of 2,6-diaminopurine for colliding with this group and keto group of thymine at position-6 by steric hindrance, and synthesized to design 2-amino-6-(N,N-dimethylamino) purine (hereinafter a base of which is designated as 1) of the formulae:

[0027]

[Formula 1]



[0028]

the base X designed as such does not form base pairing with thymine, but the base such as analogous pyridine-2-one (base X) in which oxo at position-6 of thymine is replaced by hydrogen is able to form base pair with X1 (refer to Fig. 1 and b) Consequently, formation of selective nucleic base pair of X-Y has detected (Japanese Patent Application No. 11-201450/1999).

Although formation of base pair of natural base thymine (or uridine) (refer to Fig. 1 b) could be excluded, simultaneously low rate of incorporation

of dyTP by Klenow fragment was observed due to disadvantageous effect on stacking between bases as well as insufficient suppression of incorporation of thymidine triphosphate (dTTP) to dx1.

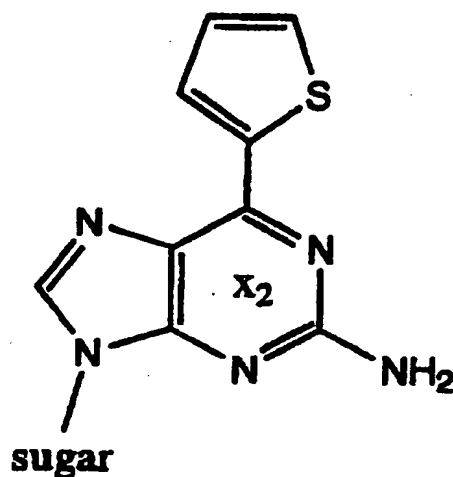
[0029]

Consequently, attempts were performed to incorporate aromatic substituents, which have no deteriorating effect on stacking between bases, to position-6 in dx as the replacement of dimethylamino group.

In the experiments, an example of incorporation of thiophene at position-6 is illustrated as follows.

[0030]

[Formula 2]



[0031]

2-amino-6-(2-thienyl)-9-(2-deoxy- β -D-ribofuranosyl) purine [dx 2: the new base of which is designated as X2, and the previously synthesized 2-amino-6-(N,N-dimethylamino)-9-(2-deoxy- β -D-ribofuranosyl) purine is designated as dx1] was synthesized. Incorporation of dyTP and ryTP for templates including this base was examined.

[0032]

Outline of preparation of amidite reagent for dx for synthesis of template DNA is illustrated in Fig. 2. In detail, refer to example 1. This amidite reagent could show similar coupling rate as same as of the commercially available amidite reagent.

[0033]

Using Klenow fragment (exo⁺), incorporation of dyTP on dx2 in the template (refer to example 2).

Following base sequences were used as template and primer.

Primer 5' - ³²pACTCACTATAGGGAGGAAGA-

Template 3' -TATTATGCTGAGTGATATCCCTCCTTCT-N-TCTCGT

In the template, a position indicated by N was bound with base X1 or base X2 (for experimental) or base A (for control), and incorporation experiments of various bases were conducted. Results are shown in Fig. 3, in which lanes 1 and 2 indicate control experiments for incorporation of cytosine (C) and thymine (T) on adenine (A).

Results indicate that rate of incorporation of dyTP on dx1 is 21% (lane 3), and that on dx2 is increased up to 40% (lane 8). As compared with the incorporation rate of 57% in the case of dTTP on natural type dA under the same condition, some improvement of incorporation was observed by using dx2. Although comparing with the case of dx1, incorporation of dCTP is increased by using dx2 (22%) (lane 11), it is not so improved value as compared with the incorporation of dyTP (40%).

[0034]

As a result of experiment using Klenow fragment (exo⁺), incorporation

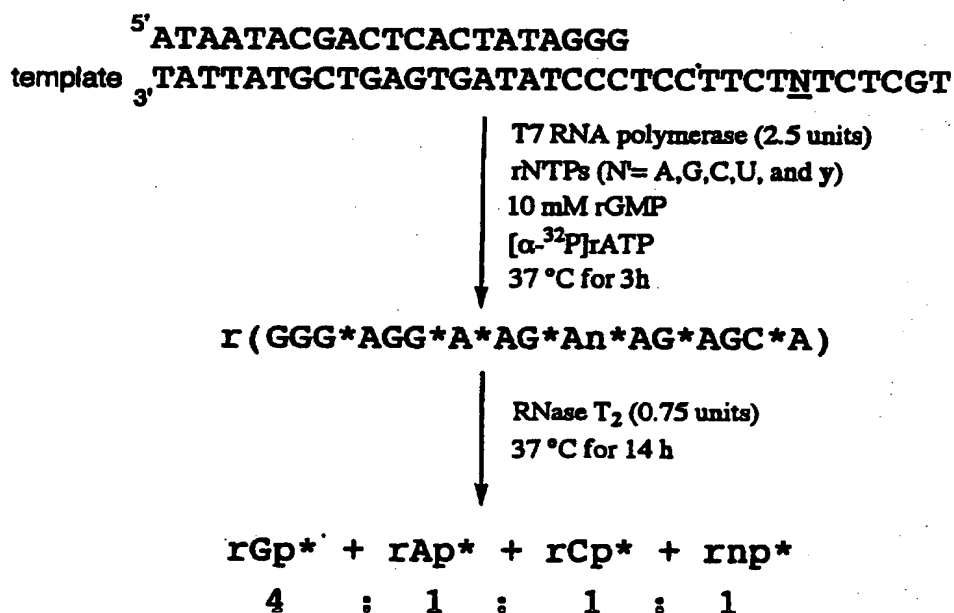
rate of dCTP on dx2 was increased by using dx2 in place of dx1. This might be due to interaction of 4-amino group of cytosine and sulfur atom in thiophene in dx2 (refer to Fig. 1 f). As is the case that sulfur atom in thiophene of dx2 is directed to a plane of base pair, electrostatic repulsion against 4-keto group of thymine (T) will be expected. This indicates that electrostatic repulsive force (refer to Fig. 1 e) can be used in addition to steric hindrance as a factor for hinder formation of base pairing. Accordingly, in thiophene in dx2, sulfur atom side might be directed to the plane of base pair.

[0035]

Incorporation of ryTP into RNA on dx2 in the template by T7 RNA polymerase was examined according to the reaction (example 3):

[0036]

[formula 3]



[0037]

RNA having a sequence of the formula:

GGG*AGG*A*AGAn*AG*AGC*A

wherein n is a base corresponding to a base N, asterisk of the right shoulder means labeling, is digested by RNase T2, then ratio of each nucleotide was calculated by 2-diemntion TLC (cellulose resin).

In Fig. 4, development of TLC is shown. Ratio of composition of each nucleotide is shown in Table 1 hereinbelow.

[0038]

[Table 1]

| template | rGp* | rAp* | rCp* | rUp* | ryp* |
|---------------------------|----------|----------|----------|----------|-----------------|
| <u>N</u> = x ₂ | 3.982(4) | 1.052(1) | 0.950(1) | 0.047(0) | 0.969(1) |
| <u>N</u> = A | 3.939(4) | 1.035(1) | 0.995(1) | 1.032(1) | not detected(0) |

Numerical values in parenthesis indicate theoretical value.

[0039]

A parenthesis in Table 1 indicates theoretical value.

Result indicates that using dx2, ryTP can be incorporated against dx2 with high selectivity. Good result has been obtained in case of using dx1 as a template in the previous experiment, and in case of using dx2 as a template to perform transcription in the similar condition, result of analysis on nucleotide incorporated in RNA on dx2 indicated that the similar high selectivity of incorporation of ryTP on dx2 was obtained.

[0040]

As explained hereinabove, it is revealed that selective formation of base pair which has never achieved in the heretofore reported artificial base pair is feasible by utilizing steric hindrance. Such artificial nucleic acid base pair can be applied to duplication and transcription of nucleic acid and to

protein synthesis system or functional nucleic acid.

[0041]

[Examples]

Following examples illustrate the present invention but are not construed as limiting the present invention.

[0042]

Example 1:

Synthesis of
2-isobutyrylamino-6-(2-thienyl)-9-[2-deoxy-3-O-[diisopropylamino]-(2-cyanoethoxy)] phosphino-5-O-dimethoxytrityl- β -D-ribofuranosyl] purine (6)
(Synthetic route is shown in Fig. 2)

[0043]

(A) Synthesis of 2-isobutyrylamino-6-iodo-9-(2-deoxy-3,5-di-O-isobutyryl- β -D- ribofuranosyl] purine (2)

2-isobutyrylamino-6-amino-9-(2-deoxy-3,5-di-O-isobutyryl- β -D-ribofuranosyl] purine (1) [Babara L. Gaffney, Luis A. Marky and Roger A. Jones, Tetrahedron, 40, 3-13 (1984)] 2.38 g (5 mmol) was heated at 60°C under argon atmosphere. n-pentylnitrite 13.5 ml (0.10 mol) and diiodomethane 25 ml (0.31 mol) were rapidly added and suspended. The mixture was irradiated by visible light using 200 W halogen tungsten lump at the distance from light source 2 cm for 3 hours under well stirring at 60°C. To the reaction mixture, saturated aqueous sodium sulfite 30 ml was added and stirred at room temperature for 3 hours. Thereafter, saturated aqueous sodium sulfite 120 ml and chloroform 150 ml were added to separate the layers. The aqueous layer was extracted twice with chloroform. The thus

obtained organic layer was dried with anhydrous magnesium sulfate and concentrated. The residue was purified by short column (developer: ethyl acetate : dichloromethane = 1 : 4) to obtain the product (2) 1.01 g (1.72 mmol) (34.4%).

[0044]

$^1\text{H-NMR}$ (270 MHz, CDCl_3) δ :

8.19 (s, 1H), 8.14 (bs, 1H), 6.42 (dd, $J = 7.4, 6.4$ Hz, 1H), 5.44 (m, 1H), 4.41 (m, 2H), 4.34 (m, 1H), 3.00 (m, 1H), 2.80 (m, 1H), 2.58 (m, 3H), 1.17 (m, 18H).

[0045]

(B) Synthesis of
2-isobutyrylamino-6-(2-thienyl)-9-(2-deoxy-3,5-di-O-isobutyryl- β -D-ribofuranosyl) purine (3)

The compound (2) 294 mg (0.5 mmol) obtained in the above (A) was dissolved in thiophene 80 ml under argon atmosphere and the solution was transferred into the photochemical reaction vessel (Pyrex). Ultraviolet ray was irradiated using 400 W mercury lamp for 24 hours under argon atmosphere. The reaction mixture after irradiation was concentrated, and the residue was purified using short column (developer: isopropanol : dichloromethane = 3 : 197) to obtain the product (3) 212 mg (0.39 mmol) (78.0%).

$^1\text{H-NMR}$ (270 MHz, CDCl_3) δ :

8.63 (dd, $J = 3.8, 1.2$ Hz, 1H), 8.17 (s, 1H), 8.10 (bs, 1H), 7.64 (m, 1H), 7.25 (m, 1H), 6.47 (dd, $J = 7.9, 1.8$ Hz, 1H), 5.44 (m, 1H), 4.43 (m, 2H), 4.37 (m, 1H), 3.18 (m, 1H), 3.00 (m, 1H), 2.61 (m, 3H), 1.24 (m, 18H).

[0046]

(C) Synthesis of 2-isobutyrylamino-6-(2-thienyl)-9-(2-deoxy- β -D-ribofuranosyl) purine (4)

The compound (3) 212 mg (0.39 mmol) obtained by the above (B) was dissolved in 1 M sodium hydroxide solution (pyridine - methanol - water = 13 : 6 : 1) 1.95 ml under ice-cooling and stirred for 15 minutes. The reaction mixture was neutralized by adding aqueous 5% ammonium chloride. Further added 1.2 g Celite to the mixture and the solvent was removed completely under reduced pressure. The residue was purified by short column (developer: 5 - 7% ethanol - dichloromethane) to obtain the product (4) 147 mg (0.37 mmol) (93.6%).

$^1\text{H-NMR}$ (270 MHz, DMSO-d_6) δ :

10.45 (bs, 1H), 8.69 (s, 1H), 8.60 (d, $J = 3.5$ Hz, 1H), 7.90 (d, $J = 4.6$ Hz, 1H), 7.32 (dd, $J = 4.6, 3.5$ Hz, 1H), 6.39 (t, $J = 6.6$ Hz, 1H), 5.34 (d, $J = 3.8$ Hz, 1H), 4.91 (t, $J = 5.3$ Hz, 1H), 4.44 (m, 1H), 3.55 (m, 2H), 2.96 (m, 1H), 2.74 (m, 1H), 2.33 (m, 1H), 1.11 (m, 6H).

[0047]

(D) Synthesis of 2-isobutyrylamino-6-(2-thienyl)-9-(2-deoxy-5-O-dimethoxytrityl- β -D-ribofuranosyl) purine (5)

The compound (4) 98 mg (0.24 mmol) obtained in the above (C) was azeotropically distilled three times with anhydrous pyridine 1 ml. The residue was dissolved in anhydrous pyridine 2 ml, added triethylamine 35 ml, dimethylaminopyridine 1.4 mg and dimethoxytrityl chloride 85 mg were added thereto and stirred at room temperature for overnight. Ethyl acetate 25 ml was added to the reaction mixture. The mixture was treated with water 25

triethylamine containing ethyl acetate. The organic layer was collected, dried with anhydrous sodium sulfate and concentrated in vacuo. The residue was purified by using short column (developer: 3% triethylamine - 32% ethyl acetate - 65% hexane) to obtain the product (6) 139 mg (0.16 mmol) (92.2%).

[0049]

^1H -NMR (270 MHz, CDCl_3) δ : 8.64 (m, 1H), 8.16 (m, 1H), 7.86 (m, 1H), 7.61 (m, 1H), 7.26 (m, 2H), 7.24 (m, 8H), 6.78 (m, 4H), 6.45 (m, 1H), 4.75 (m, 1H), 4.23 (m, 1H), 3.75 (m, 6H), 3.70 (m, 4H), 3.36 (m, 2H), 2.75 (m, 2H), 2.62 (m, 1H), 2.48 (m, 1H), 1.95 (m, 1H), 1.18 (m, 18H).

^{31}P -NMR (270 MHz, CDCl_3): 149.51, 148.43 ppm.

[0050]

Example 2:

Single nucleotide insertion reaction using Klenow fragment (exo⁺)

A solution containing [5'- ^{32}P] labeled primer DNA (20-mer, 4 mM), template DNA (35-mer, 4 mM) and 2x Klenow fragment buffer (TAKARA) were annealed at 95°C for 3 minutes, 40°C for 3 minutes and 4°C for 7 minutes. A solution of equimolar amount of 40 mM dNTP and Klenow fragment (exo⁺) (2 unit/ml, For Sequencing, TAKARA) were added thereto and incubated at 37°C for 30 minutes. Equimolar amount of 10 M urea BPB dye solution was added and kept at 75°C for 5 minutes and electrophoresed with 20% polyacrylamide - 7M urea gel. Products were analyzed by using Phosphoroimager plate. Result is shown in Fig. 3.

[0051]

Example 3:

Transcription by T7 RNA polymerase

A solution containing template DNA 1mM, in which promoter region has duplicated strands, T7 RNA polymerase 2.5 units, 2mM rNTP, and [α - 32 P] rATP 0.1mCi/ml [40 mM Tris-HCl (pH 8.0), 8 mM MgCl₂, 2 mM spermidine, 5 mM DTT, 0.01% Triton X-100, 10 mM rGMP] were prepared and incubated for 3 hours. 10 M urea dye was added and kept at 75°C for 3 minutes to terminate the reaction. The product RNA (16-mer) in this solution was purified by using electrophoresis with 20% polyacrylamide gel. RNA was digested by 0.75 units RNase T2. Ratio of each nucleotide was determined by 2-dimension TLC (cellulose resin). In Fig. 4, result of development of TLC is shown. Ratio of each nucleotide is shown in Table 1.

[0052]

Example 4:

Synthesis of primer and template

Following primer and template were synthesized conventionally by using DNA/RNA synthesizer Type 392, The Perkin-Elmer, Applied Biosystems Div., and cyanoethylamidide reagents of dA, dC, dG and dT, which were available from The Perkin-Elmer, and dX₂ of cyanoethylamidite reagent in Example 1.

Proviso that removal of protective group for 2-amino group of dX₂, i.e. isobutyryl group, could not completely be performed under usual basic condition after synthesis of oligomer (conc. ammonia at 55°C for 10 hours), consequently, treatment under the condition at 80°C with conc. ammonia for 10 hours was performed.

[0053]

[Effect of the Invention]

The present invention indicates that selective base pair formation which could never be achieved by the heretofore reported artificial base pair, can be realized by utilizing steric hindrance and electrostatic repulsion as well as stacking action. By utilizing the method of the present invention, artificial nucleic acid base pair of the present invention can be applied on replication and transcription of nucleic acid, and protein synthetic system or functional nucleic acid.

[Brief Description of Drawings]

[Fig. 1] Fig. 1 shows conventional base pair (fig.1 a and b) and novel artificial nucleic acid base pair (X2-Y)

[Fig. 2] Fig. 2 shows synthetic scheme for amidite reagent for dX2 of nucleic acid having base X2 of the present invention.

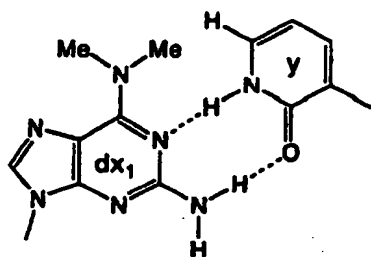
[Fig. 3] Fig. 3 shows incorporation of various base against X2 of the present invention. Drawing is replaced by photograph.

[Fig. 4] Fig. 4: RNA which generated when all rNTP was coexisted was purified by electrophoresis, then digested with RNase T2 and analyzed by 2-dimension TLC. Fig. 4 shows its analysis result.

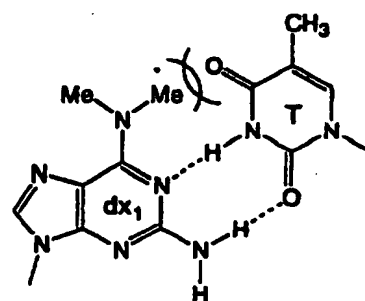
[Document] Drawings

[Fig. 1]

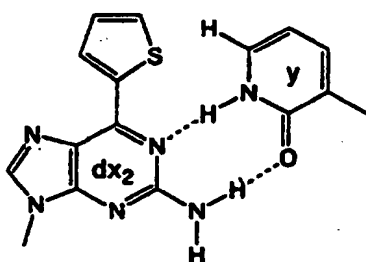
a



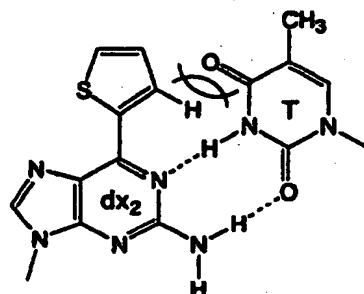
b



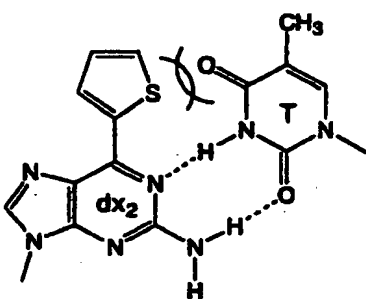
c



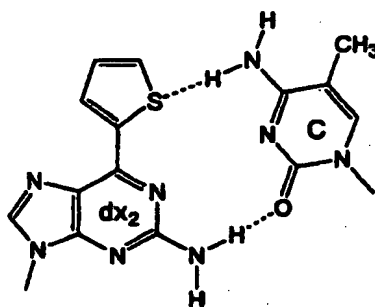
d



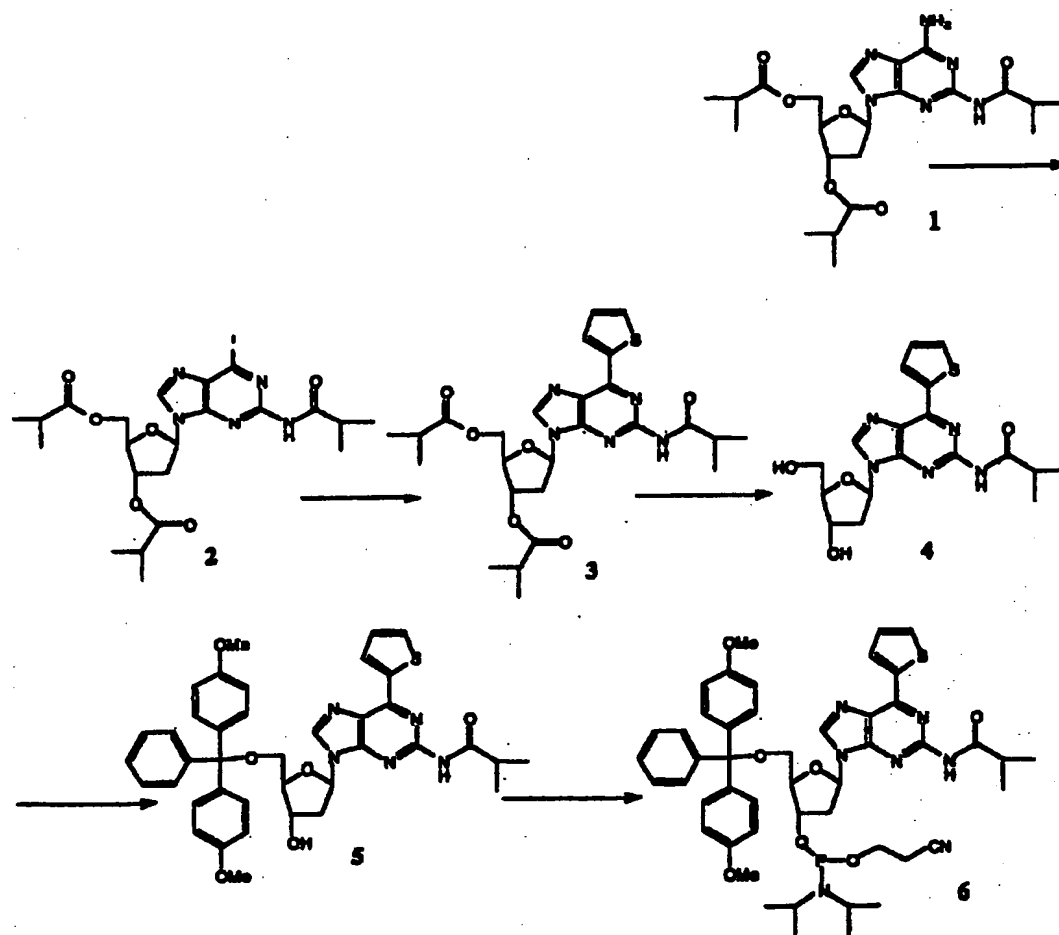
e



f



[Fig. 2]

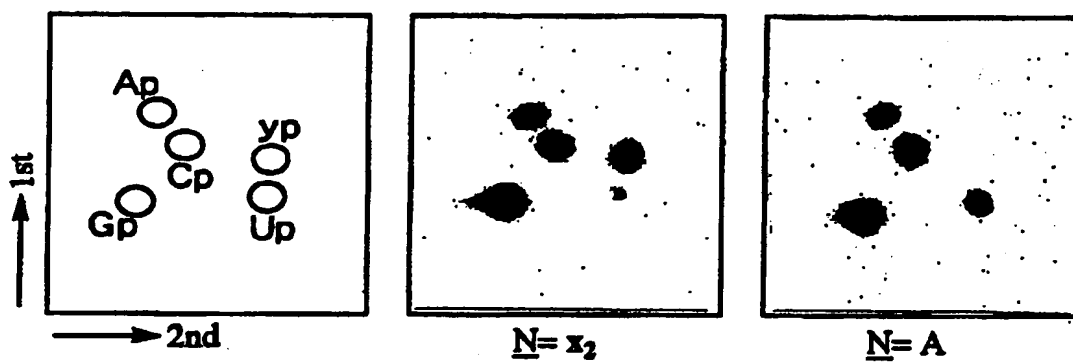


[Fig. 3]

| | | | | | | | | | | | | |
|---------------------|---|----|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| template : <u>N</u> | A | A | X ₁ | X ₁ | X ₁ | X ₁ | X ₁ | X ₂ | X ₂ | X ₂ | X ₂ | X ₂ |
| dNTP | C | T | Y | A | G | C | T | Y | A | G | C | T |
| efficiency (%) | 8 | 57 | 21 | 10 | 9 | 13 | 12 | 40 | 9 | 8 | 22 | 12 |

| | | | | | | | | | | | | |
|--------|---|---|---|---|---|---|---|---|---|----|----|----|
| 21-mer | | | | | | | | | | | | |
| 20-mer | | | | | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |

[Fig. 4]



[Document] Abstract

[Abstract]

[Problem]

The present invention provides ideas for selective formation of novel artificial nucleic acid base pair as a result of recognition of base pairing by polymerase such as DNA polymerase by utilizing steric hindrance between base pairs, further generating steric hindrance between only base pair plane without giving deterioration for stacking between bases and more preferably by selecting bases utilizing electrostatic repulsion against natural bases.

[Means for Solving the Problems]

The present invention relates to a method for constructing selective base pair comprising introducing a group having ability to form steric hindrance, electrostatic repulsion and stacking action in nucleic acid base. Further, the present invention pertains the novel artificial nucleic acid having properties to form selective base pair by applying steric hindrance and stacking action, a process for production thereof, codon containing the same, nucleic acid molecule containing the same, a process for production of novel protein using the above nucleic acid molecules or non-natural gene.

[Chosen Drawing] None